Diachronic regularities explaining the tendency towards explicit analytic marking in Balkan syntax

Lindstedt, Jouko

De Gruyter Mouton
2018

Lindstedt , J 2018 , Diachronic regularities explaining the tendency towards explicit analytic marking in Balkan syntax . in I Krapova & B Joseph (eds) , Balkan syntax and (universal) principles of grammar . Trends in Linguistics. Studies and Monographs , no. 285 , De Gruyter Mouton , Berlin , pp. 70-84 . https://doi.org/10.1515/9783110375930-005

http://hdl.handle.net/10138/309437
https://doi.org/10.1515/9783110375930-005

unspecified
acceptedVersion

Preprint. To appear in Brian Joseph & Iliana Krapova (eds.), *Balkan Syntax and Universal Principles of Grammar.* Mouton de Gruyter.

Diachronic regularities explaining the tendency towards explicit analytic marking in Balkan syntax

Jouko Lindstedt (University of Helsinki)

A growing number of linguistic studies argue that contact among languages is reflected in their structure in particular ways and that certain types of contact situations simplify grammar, whereas other types of contact situations, and especially the relative isolation of a language, may complexify it (Kusters 2003; McWhorter 2001, 2011; Trudgill 2002, 2011; Dahl 2004: 280–285). In this paper I will argue that the changes that made the languages of the Balkan linguistic area converge structurally were neither clearly simplifying nor clearly complexifying. These changes represent a tendency towards a certain syntactic type, that is, explicit analytic marking, whose rise can be explained by diachronic regularities that are partly structural, partly sociolinguistic in nature. I will also argue that the Balkans represent a third type of contact situation besides the two types distinguished by Trudgill (2011).

1. Simplification, complexification and the Balkan languages

The structural convergence in the Balkan linguistic area (Sprachbund) has been used as an example of both contact-induced simplification and contact-induced complexification (for lists of the relevant structural features, the so-called Balkanisms, see Joseph 2013 and Lindstedt 2000a).

Hinrichs (2004) emphasizes similarities between the Balkan languages and creoles, though he admits that there are also differences. Hinrichs sees the changes that brought about the Balkan linguistic area as examples of "unnatural" change owing to an extreme contact situation (Hinrichs 2004: 142–144); this comes close to McWhorter's (2011) idea of creoles as languages whose normal accumulation of complexity has been interrupted or Dahl's (2004: 281) concept of "suboptimal transmission". As characteristics of the Balkan Sprachbund, Hinrichs (2004: 170) mentions "Kontaktzwänge, Reduktionismus, Vereinfachung, Rekonstruktion, Kreolisierung und weiteren Ausbau". As the sociolinguistic context of this Balkanization of languages, he sees the orally orientated ("oral geprägte") culture of the

Balkans, which he compares to the "westafrikanischen Kulturen der vermuteten Substratsprachen der KS" [=Kreolsprachen] (Hinrichs 2004: 163–167).

As for this sociolinguistic explanation, I find Hinrichs's proposal unconvincing. On the one hand, most changes in most languages of the world have always originated in their spoken varieties, and it is difficult to see why the Balkans (or West Africa, for that matter) would have been special in this respect. On the other hand, Southeast Europe was the home of two of the three most important written languages of mediaeval Europe – Greek and Old Church Slavonic – and it is therefore difficult to see what would make the Balkans a region with a distinctively "oral" culture by comparison with other regions in Europe.

But it is also difficult to agree with Hinrichs's characterization of Balkanization as simplification. For instance, the verb systems of the Balkan languages resemble each other, but they are much more complex than those of any creole language; a modern grammar of standard Albanian (Demiraj, ed., 2002) distinguishes ten tenses and six moods. Case inflection has been reduced in all Balkan languages to some extent, but only Balkan Slavic has lost case marking completely. What should also be borne in mind is that Balkan Romance is the sole branch of Romance that has *preserved* case distinctions in nouns. And the appearance of enclitic definite articles in several Balkan languages did not simplify those languages, at least for the speaker, because it required obligatory coding of a feature that initially did not have explicit marking. Pidgins and creoles often do not make use of definite articles even when their lexifier languages possess them (Bruyn 1995: 259).

A definite article or the clitic doubling of the object, to give another Balkan example, can, of course, be seen as increasing the *redundancy* of the message to the benefit of the hearer, and Hinrichs does mention this characteristic. Redundancy has also been mentioned by other scholars in connection with Balkanisms (e.g. Hauge 1977). However, Hinrichs seems to think that redundancy is naturally linked to simplification and creolization, whereas Dahl (2004: 5–17) and Trudgill (2011: 62) consider redundancy to increase complexity. Dahl (2004: 9) even defines redundancy with the help of complexity: "A message is redundant if there is a less complex message that could – in principle – transfer the same amount of information, that is, if more communicative resources are spent on it than are theoretically necessary for its successful delivery".[1]

[1] I assume that the hedges "in principle" and "theoretically" are necessary for natural language since once a redundant marker has been fully grammaticalized, there may not be a less redundant grammatical way to express the message in the same synchronic state of the language in question. For instance, an agreement marker may be obligatory in a given language, but in comparing it with other languages, we see that it is not necessary for a functioning human language.

As redundancy and complexity are connected in this fashion, it is natural that in his book *Sociolinguistic typology,* Peter Trudgill (2011: 34, 42) uses Sprachbünde or linguistic areas as examples of the type of contact situation that leads to complexification; his conclusion is thus diametrically opposed to Hinrichs's. Trudgill does not deal with the Balkans in greater detail in his book, but in an earlier article (Trudgill 2002: 710–711) he presents the Balkan loss of the infinitive as an instance of complexification and increasing redundancy. He compares the Greek sentence

(1)	*θél-o*	*na*	*γráp-s-o*
	want-PRS.SG1	COMP	write-PFV-PRS.SG1
	'I want to write'		

with the corresponding English sentence *I want to write* and notes that the information about the first person singular is marked in English only once, but in Greek twice and therefore in a more redundant fashion. But notice that the subject of the embedded verb could also be different from that of the matrix clause. When this is taken into account, the Balkan pattern is actually more symmetrical than the English system (I use Macedonian here to represent the Balkan system):

(2)	*sakam da dojdam*	*– I want Ø to come*
	sakam da dojdeš	*– I want you to come*
	sakaš da dojdam	*– you want me to come*
	sakaš da dojdeš	*– you want Ø to come*

In English, there are syntactic rules defining when the embedded verb can or must appear in the infinitive and when its subject can or must be dropped. In Macedonian and other Balkan languages, the underlying finite forms and their person markings are preserved. Balkan grammar is here rather simpler, not more complicated, than that of most European languages. It can in fact be argued that in the Greek *θélo na γrápso* 'I want to write' or the Macedonian *sakam da dojdam* 'I want to come', there is no real redundancy in the person marking at all, since the first person singular is opposed to all other persons, not only in the matrix verb but also in the embedded verb, independent of the matrix. With clearly modal verbs indicating possibility ('can', 'may') or necessity ('must'), such a distinction between the matrix person and the embedded person is, of course, not valid. But exactly these verbs are often used in the

Balkan languages in the unmarked third person singular only, so that it is the lexical verb that shows the person without redundancy.

Notice also that the Balkan loss of the infinitive means that the verb has one inflectional category less, which certainly can count as a paradigmatic simplification.

In general, the grammatical Balkanisms cannot all be unequivocally characterized as simplifying or as complexifying. Several of them, such as the reduction of the case system, recipient/possessor merger, goal/location merger, *relativum generale*, or analytic comparison of adjectives, do decrease the obligatory paradigmatic choices to be made and are in that sense simplifying, but part of the same information must then be encoded syntagmatically. The rise of the analytic *volo* future meant a similar syntagma-for-paradigm swap. Enclitic articles, clitic doubling and other analytic means of marking the arguments increase redundancy and are in that sense complexifying. Grammaticalized evidentials or the *habeo* perfect are new verb categories and most clearly complexifying.

However, the essential typological characteristics of the Balkan linguistic area cannot be described along the simple/complex axis at all. What is typical of most Balkanisms is the *explicit analytic marking* of syntactic relations and other grammatical categories by prepositions, pronominal clitics, articles, particles and other function words. Explicit analytic marking can be typologically opposed to inflectional (synthetic) marking, as well as to implicit analytic marking with word order alone and to the absence of marking.

What factors have brought about such a structural type? I will propose three diachronic regularities to explain the mechanism of change towards explicit analytic marking in the Balkan languages. My proposals should be considered conjectures, that is to say, informed hypotheses that are not incongruent with what we know about the historical changes in the Balkan languages, their past sociolinguistics and language contacts in general. I hope they will give ideas for further research.

2. Borrowing increases analytism

The convergence of the Balkan languages presupposes borrowing. For most Balkanisms, a single source language cannot be indicated, to be sure, but even when structural Balkanisms arose through mutual reinforcement of change among languages (Lindstedt 2000a), having thus multiple sources, the patterns and constructions must have been copied back and forth between languages.

Now, it has long been known that analytic constructions are more likely to be borrowed than inflectional categories. Weinreich (1970 [1953]: 41) wrote: "Significantly, in the interference of two grammatical patterns it is ordinarily the one which uses relatively free and invariant morphemes in its paradigm – one might say, the more explicit pattern – which serves as the model of imitation". Thomason (2001: 69) writes that "less tightly structured features are easier to borrow than features that fit into tightly integrated closed structures", inflectional morphology being an example of the latter. According to Dahl (2004: 127–128), "what is borrowed, or calqued (i.e. translated), in grammar will most frequently be periphrastic constructions or free markers, and less often affixes, although the latter is also observed to happen". He proposes that structures more likely to be borrowed are less "mature"; by mature structures he means a structure that exists in a language only when it has passed through a specific earlier stage (along a grammaticalization path, for instance).

All this is quite uncontroversial, but I surmise that there is an important corollary: the more a language borrows from other languages (beyond mere lexical items), the more analytic it is bound to grow in the long run. If analytic constructions always have a higher probability of entering a language than synthetic constructions do, the former will gradually become more frequent, and this will happen more rapidly if there is large-scale borrowing. This means that an important structural feature of the Balkan languages, their analytism, is a result of those social circumstances that enhanced borrowing among those languages. Of course, these social circumstances in their turn have to be explored and explained, but at any rate, such a simple connection between social context and language structure seems to exist.

Notice that this conjecture does not require the source language of the borrowings to be more analytic than the target language; what is needed is only the right kind of social context for intense borrowing in general. It is the borrowing language that selects the analytic constructions in the source language. Of course, if there are none – if the source language is a polysynthetic language, for instance – the situation is different, but I assume that in such circumstances grammatical borrowing would be rare anyway.

Moreover, the borrowing language may restructure the borrowed pattern so that it becomes more analytic than it was in the source language. To take an example outside the Balkans, Finnish Romani has borrowed from Finnish the periphrastic perfect and pluperfect tenses, which are now opposed to the simple past tense of the type *mekjas* '(s)he left / was leaving', *rakkadas* '(s)he spoke / was speaking'. The Finnish perfect and pluperfect are formed with the present or past auxiliary 'to be', respectively, and the past participle of the verb; being periphrastic constructions, they are likely to be borrowed. However, Finnish

Romani has only borrowed the use of the auxiliary 'to be'; the main verb has not been changed into a participle. The Finnish Romani perfects are of the type *hin mekjas* '(s)he has left', *hin rakkadas* '(s)he has spoken', while the pluperfects are of the type *sas mekjas* '(s)he had left', *sas rakkadas* '(s)he had spoken', where both the auxiliary and the main verb are finite; thus, *sas rakkadas* is literally "was spoke". The Finnish Romani periphrastic verb constructions are more analytic than their Finnish models because the auxiliary does not govern the inflectional form of the main verb, but is simply concatenated with it. As a further development, the third-person auxiliary is further generalized optionally for all the persons. The first person singular *som mekjom* 'I have left' is then replaced with *hin mekjom,* where only the main verb shows the person; this is a further move towards analyticity because the tense has received an invariant marker (the examples are taken from Granqvist 2011: 117–124).

All of this means that if an areal group of languages begins to borrow structures from each other, as happened in the Balkans, the degree of analytism in each of them may in the long run become higher than any one language among them had in the beginning. In the Balkans we also see geographically that language varieties spoken in areas of intensive and multilateral contact are more analytic. The contact area of Greek, Albanian, Macedonian, Aromanian, Romani and Turkish in the Central Balkans, around the lakes Ohrid and Prespa and south of them, is the home of the most Balkanized varieties of these languages, as pointed out by Asenova (2002: 17); she sees this centre of Balkanization to be approximately delimited by the river valleys of Shkumbin, Vjosë/Aóos and Vardar/Axiós. This is also the sphere of influence of the western part of the ancient Via Egnatia.

The regularity "borrowing increases analytism" proposed here would not be universally true if analytic structures were also generally lost through obsolescence more quickly than synthetic structures. But what we know about language change certainly does not render this plausible. For instance, in several European languages periphrastic perfects are becoming or have become generalized past tenses or perfective pasts (Lindstedt 2000b: 365–374); we would not expect old synthetic past tenses to recapture the semantic field they have already lost to the new perfects.

The only situation in which this regularity would perhaps not be accurate is when widespread childhood bilingualism is involved, because the preference for analytic structures may be linked with adult second-language learning (Trudgill 2011: 40–43). But, unlike Trudgill, I do not believe that the Balkans ever belonged to this type, as I will discuss in Section 5 below.

As for the simplification / complexification dimension, borrowed analytic structures simplify the language if borrowing is replacive, i.e. if the new structures supersede old, possibly synthetic structures. But if borrowing is additive, it may introduce new distinctions and thus complexify the grammar. The borrowing of grammatical evidential distinctions from Turkish into Balkan Slavic created a new grammatical category and thus made the structure of the borrowing languages more complex, even though evidentiality is expressed in them by periphrastic verb forms. The borrowing of the *habeo* perfect from Greek and Balkan Romance into Macedonian dialects, in its turn, was first simply additive, but when it caused the old *esse* perfect to specialize for evidential uses in a growing number of dialects (Graves 2000; Lindstedt 2000b; Bužarovska & Mitkovska 2010), the result was simplification in the sense of greater morphological transparency ("one meaning, one form", Anttila 1972: 100–102; cf. Trudgill 2011: 21).

3. Convergence of different structures increases analytism

The tendency towards explicit analytic marking is manifested in the nominal system of the Balkan languages in the decrease of case distinctions and the increased use of prepositions, although only Balkan Slavic and dialects of Aromanian have completely lost the cases with the exception of the personal and some other pronouns. (The vocative is left out of this discussion as it is not a syntactic case that marks the argument structure of the sentence.) In the linguistic geography of Europe, the Balkans are not unique in this respect; rather they form an intermediary zone between the caseless West and the case-preserving East, as is also shown by comparing Romanian with the other Romance languages. The complete loss of case in Balkan Slavic was a complicated process influenced by several factors (Wahlström 2015).

The remaining case systems in Albanian, Greek and Balkan Romance are remarkably similar, with a genitive-dative case that merges the marking of the possessor and the recipient functions. But another side of the same coin, less often mentioned, is the isosemantism of the most common prepositions, i.e. the fact that their meaning and use is so similar that if one knows the preposition used in one Balkan language, one can usually correctly predict the preposition used in another (Asenova 2002: 97–104).[2] As in other typical Balkanisms, it is the meanings, not the forms of the prepositions that have converged. In Balkan Romance and dialects of Macedonian, a preposition is used even to mark direct objects, especially those

[2] A notable exception is the Greek preposition *se,* which in its locative use corresponds to several distinct prepositions in other Balkan languages.

referring to humans.[3] Clitic doubling of the direct and indirect objects is also part of the same picture: the marking of the argument structure of the sentence has moved in an analytic direction, and the markers exhibit a high degree of intertranslatability[4] among the languages, especially among their coterritorial dialects.

Compared with the nominal system, the Balkan verbal system is remarkably complex, with a wide variety of grammaticalized aspectual, temporal, modal and evidential distinctions, part of which are still expressed inflectionally. The oppositions among the past tenses in the Balkan languages are similar, as is the expression of counterfactual conditions (Asenova 2011: 220–274; Gołąb 1964; Lindstedt 2002). But there are convergent analytic forms too, notably the structure "complementizer + finite verb" that is used instead of the infinitive (as discussed in Section 1) and the use of a future auxiliary that has lost part of its verbal inflection or has become an uninflected particle. Their role in the Balkan verbal system is similar to that of the prepositions in the nominal system.

I surmise that the reason the Balkan verbal system has remained more complex than the nominal system is that the common Indo-European inheritance of the Balkan languages already contained similar distinctions: there were aspectually opposed past tenses, there were past participles that could be used to form new periphrastic forms in the same fashion, and there were the copula 'to be' and the transitive possessive verb 'to have' that could be used as auxiliaries in new periphrastic perfects in languages that did not already possess such forms. The different case systems, by contrast, could not be harmonized in the same fashion before they were reduced to a tripartite system of nominative : accusative : genitive-dative, and much of their functional load was shifted to prepositions exhibiting isosemantism among the languages.

I propose the following diachronic regularity: if the corresponding subsystems of two or more languages tend to converge, then the resulting subsystems will be more analytic if they were structurally different to begin with, whereas relatively similar subsystems may reach convergence on a lower level of analyticity.

The regularity proposed does not say anything about why the languages must converge in the first place; some of the sociolinguistic mechanisms behind this will be discussed in Sections 4 and 5 below. Nor does this regularity directly explain why Balkan Slavic had to go so far as to lose its case distinctions altogether, although it had already developed a system

[3] The preposition used is *na* in Macedonian and *pe, pi* in Balkan Romance; see, for instance, Markovik' (2007: 91–92) for the Macedonian and Aromanian dialects of the Ohrid-Struga region.
[4] This notion seems to have been introduced to contact linguistics by Gumperz and Wilson (1971).

similar to that in Albanian, Greek and Balkan Romance (Wahlström 2015: 81–85; Gołąb 1997). As shown by Wahlström, the reasons for this further development in Balkan Slavic were certainly manifold, but in the context of this article it can be pointed out that with analytic argument marking, in other words, with prepositions and clitic doubling, an even higher grade of intertranslatability among languages could be attained than with case endings. Exactly why speakers of Balkan Slavic would have striven for high intertranslatability will also be discussed in Section 4 below.

On the other hand, the Balkan convergence should not be exaggerated. Albanian and Balkan Slavic, for instance, have both borrowed the category of evidentiality from Turkish, but despite many similarities between them in this respect, there are also substantial differences in the form, meaning and use of the indirect evidentials (Friedman 1986; Makartsev 2014). The regularity proposed in this section does not suggest that strong convergence and intertranslatability are inevitable in a particular kind of contact situation; it only says that *if* convergence is to be attained, then the more different the initial subsystems were, the more analytic structures are required to replace them.

4. Both L2 speakers and bilingual L1 speakers count

The greatest challenge in Balkan linguistics is to anchor the explanations for the convergent development of the languages in their past sociolinguistic context, a context that is known to us only indirectly and through sources that were usually not written for the purpose of recording linguistic observations. The danger of circularity lurks: we explain the observed convergence with past bilingualism, but the convergence is also the proof that such bilingualism must have existed.

A fundamental question is *who* Balkanized the languages: was it (i) the L1 speakers who were influenced by their strong L2 languages? Was it perhaps (ii) the L2 speakers influenced by their L1 languages? Or was it, as proposed by Civ'jan (1965: 14ff., 183ff.), among other scholars, (iii) that L1 speakers were attempting to speak in structures that would be maximally comprehensible to the speakers of another Balkan language? All three explanations are to some extent plausible, but there are also problems with each: in scenario (i), what about the great number of speakers who did not need to learn any L2 well, such as the speakers of the prestige language Greek? In scenario (ii), how could L2 speakers have really initiated changes in a language that had a native-speaker majority, apart from some few

situations in which a large-scale language shift occurred? In scenario (iii), how might these bilingual encounters have changed the way anyone spoke in other, monolingual situations?

Questions of this kind are difficult to answer by observing the present-day Balkans, where mutual bilingualism has disappeared with the rise of nation states and only the bilingualism of the dwindling minorities is left. But an indirect way of approaching the problem is first to note that among all the Balkan language groups, it was Slavic that was the most affected by the tendency towards explicit analytic marking, Macedonian being the most Balkanized of all, at least on the level of standard languages. (I think this characterization of Balkan Slavic is uncontroversial at a general level, even if not everyone would agree with the way I counted Balkanization indices in Lindstedt 2000a: 232–234.) What does this fact tell us?

Macedonian is located in the geographical centre of Balkanization (see Sect. 2), which partly explains its high number of Balkanisms. But an explanation valid for all of Balkan Slavic would begin by noting that Slavic was certainly not the most prestigious language in the Balkans during the Ottoman period, but neither was it lowest on the prestige scale. Greek was the prestige language among the Christian population – to the extent that all educated users of Greek irrespective of their native language shared the same "Romaic" identity (Detrez 2008, 2015) – while Ottoman Turkish was the language of the state. On the other hand, the prestige of Slavic was certainly higher than that of Romani and, it is safe to assume, also somewhat higher than that of Albanian and "Wallachian" (Balkan Romance), whose speakers used mainly Greek and Church Slavonic as their written languages. In all, Balkanization was strongest in Slavic, near the middle of this prestige scale, and Greek and Romani, at the two opposite ends of the same scale, clearly possess a lesser number of Balkanisms (Lindstedt 2000a: 232–234).[5]

I surmise that what makes the middle of the prestige scale special is the *combined* role of L2 speakers and L1 speakers regularly using other languages. Balkan Slavic certainly had more L2 speakers than the languages below it on the prestige scale, especially Romani. In the Central Balkans the direction of the language shift has been from Romance to Slavic (Gołąb 1997). And it is reasonable to assume that the Christian and Muslim Slavs communicated with the Muslim Albanians mainly in Turkish or Slavic. On the other hand, Balkan Slavic had

[5] Even now Romani is still not often mentioned among the languages of the Balkan linguistic area; Asenova (2002: 16–19), for instance, does not bring it up in discussing the languages of which the Sprachbund consists. It is true that Romani has fewer Balkanisms than Balkan Slavic, Balkan Romance or Albanian, but no such clear line can be drawn between Romani and Greek. The neglect of Romani is thus mainly a result of the history and sociology of scholarship, not of the objective properties of this language.

more L1 speakers regularly using other languages than Greek did (or Turkish, for that matter), because the speakers of a prestige language have fewer reasons to learn other languages well.

Weigand (1895: 6) reported that the Aromanians of Monastir (present-day Bitola in Macedonia), "at least the men", knew, besides their native language, Bulgarian and Greek; "most of them" also knew Turkish and Albanian, and many understood even (Judeo-)Spanish. From this list we may conclude that Albanian was lower on the local scale of prestige (or at least utility value) than Macedonian (what Weigand called Bulgarian), though we must of course keep in mind that in those Ottoman provinces where there were more Albanian and fewer Slavic speakers the situation was certainly different.

My conjecture is that grammatical borrowing that favours change towards analytism may occur both when L1 speakers regularly use another language and also when L2 speakers transfer features from their native languages. Furthermore, the combined effect of these two bilingual speaker groups when both are large in a given language community is stronger than if only one group were large. This explains nicely why Balkan Slavic (and especially Macedonian) has developed several Balkan features more than have the other Sprachbund languages. The combined effect of L2 speakers and L1 speakers regularly using other languages also explains why Greek is not the main source of Balkanisms (*pace* Sandfeld 1930, see Lindstedt 2000a: 236–237), or why Romani has not adopted the greatest number of Balkanisms, but the sociolinguistic focus of Balkanization seems to have been around the middle of the prestige scale.

This also fits well with Joseph's (2013: 625) explanation emphasizing "intense, intimate and mutual multilingualism" in the Balkans and describing mutual accommodation in bilingual contact situations, in which L2 speakers have interference from their native languages; L1 speakers also try to use the kinds of patterns in their own language that they know have analogues in the other languages. Notice that in this model an L2 may influence a speaker's L1 not only because the L2 may be a much-used prestige language, but also because the speaker regularly uses L1 with the native speakers of L2 and has therefore been accustomed to accommodating their own L1 usage towards this L2.

This **principle of combined effect** could theoretically mean two different things: either the combined effect of the two kinds of bilingual speakers favours all kinds of explicit analytic marking or else the L1>L2 interference favours one type of feature, the L2>L1 interference another type of feature and the "combined effect" appears only insofar as these types are not distinguished from one another, but counted together. Yet in the latter instance, we would expect to find very different sets of Balkanisms at the two ends of the prestige

scale, Greek and Romani, and this does not seem to be the case. It can thus be assumed that *both* L1 speakers regularly using other languages *and* L2 speakers created explicit analytic marking in the Balkan languages, albeit in different proportions.

5. The Balkans: A third kind of contact situation

Trudgill (2011: 40–43) distinguishes two sociolinguistic types of contact situations, which have opposite effects on language complexity. Simplification occurs in contact situations that are dominated by "untutored, especially short-term, *adult* second language learning" (original emphasis; this type is also the main focus of McWhorter 2011). On the other hand, complexity increases "in long-term, co-territorial contact situations which involve childhood […] bilingualism where young speakers know two or more languages natively or at least extremely well". Trudgill repeatedly gives Sprachbünde as examples of the latter kind.

Now, I have argued that the effects of the Balkan contact situation were neither clearly simplifying nor complexifying, but rather that explicit analytic marking is the property common to most grammatical Balkanisms. This result does not fit Trudgill's two types very well, nor does the Balkan sociolinguistic context correspond to either type. In the Balkans we certainly have a "long-term, co-territorial contact situation", but there is no reason to assume that this situation included much childhood bilingualism, except for some ethnically mixed urban centres. On the other hand, adult bilingualism, especially adult male bilingualism, was not based on "short-term" second-language learning; rather it was a stable property of the linguistic situation.

In the Balkans of the Ottoman period, it was men who were responsible for most of the economic and administrative contacts among different ethnic groups, because men represented the family before the state and religious authorities and traded in products of livestock breeding or crafts. Men also participated in *gurbet,* seasonal work migrations; women could do so more rarely and only until they became married (Hristov 2008: 3). This is not to deny the possibility of female bilingualism or childhood bilingualism at various places and times, but at any rate, the Balkan sociolinguistic situation did not correspond to either of Trudgill's two types.

I propose that there were two crucial sociolinguistic factors that created the specificity of the Balkan situation. In the Ottoman period there was no single lingua franca in the region. Turkish was the state language, and Greek was the prestige language of learning among the Orthodox ("Romaic") population, but neither of them alone combined both state and religious

authority as far as the Balkan Christians were concerned. Another important factor was that languages were important symbols of group identity, and large-scale language shifts did not occur. Greek was perhaps on its way to becoming the identity language of the Balkan Christians (Detrez 2008, 2015), but this development was cut short by the new nationalist movements, which finally terminated the Sprachbund formation altogether.

The beginning of the Sprachbund formation must be placed in the last centuries of the Byzantine era,[6] and here we have far fewer facts upon which to construct sociolinguistic hypotheses. Greek was certainly the unrivalled prestige language of the mediaeval Balkans, but Church Slavonic was used as the written language of the Slavic kingdoms; at any rate, the time was much more turbulent than in the later Pax Ottomanica. For six decades during the thirteenth century, Constantinople was the capital of the Latin Empire *(Imperium Romaniae)* founded by the crusaders, and in that time the Greek language was strongly influenced by Western Romance, notably French (Horrocks 2010: 345–359), which may in fact account for some of the apparent Balkanisms in Greek.

To sum up, a third type of contact situation can be posited on the basis of Balkan facts: adult-based long-term, stable, mutual and intense multilingualism which does not lead to outright simplification or complexification, but favours such explicit analytic grammatical marking that increases direct intertranslatability among the languages.

Acknowledgements. I wish to thank the participants of the Balkan panel at the Societas Linguistica Europaea meeting in Split in 2013 and especially the anonymous reviewer of this volume for their useful comments and suggestions. Unfortunately I was able to take into consideration only part of them.

References

Anttila, Raimo. 1972. *An introducttion to historical and comparative linguistics.* New York: Macmillan.

Asenova, Petja. 2002. *Balkansko ezikoznanie: Osnovni problemi na balkanskija ezikov sǎjuz.* 2nd edn. Veliko Tărnovo: Faber.

Bruyn, Adrienne. 1995. Noun phrases. In Jacques Arends, Pieter Muysken & Norval Smith (eds.), *Pidgins and creoles: An introduction* (Creole Language Library), 259–269. Amsterdam: John Benjamins.

Bužarovska, Eleni & Liljana Mitkovska. 2010. The grammaticalization of the habere-perfect in Standard Macedonian. *Balkanistica* 23. 43–65.

Civ'jan, T. V. 1965. *Imja suščestvitel'noe v balkanskix jazykax. K strukturo-tipologičeskoj xarakteristike balkanskogo jazykovogo sojuza.* Moskva: Nauka.

[6] Joseph (2013: 618–619) puts the beginning of the Sprachbund formation roughly at 1000 CE.

Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity* (Studies in Language Companion Series 71). Amsterdam: John Benjamins.

Demiraj, Shaban (ed.). 2002. *Gramatika e gjuhës shqipe,* I: *Morfologjia.* Tiranë: Botim i Akademisë së Shkencave.

Detrez, Raymond. 2008. Ezik i literatura v "romejskata" obšnost. In Gajda, Stanisław (ed.), *Tożsamość a język w perspektywie slawistycznej,* 151–175. Opole: Uniwersytet Opolski, Instytut Filologii Polskiej.

Detrez, Rajmond. 2015. *Ne tărsjat gărci, a romei da bădat. Pravoslavnata kulturna obštnost v Osmanskata imperija. XV–XIX v.* S.l.: Kralica Mab.

Friedman, Victor. 1986. Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian. In Wallace Chafe & Johanna Nichols (eds.), *Evidentiality: The linguistic coding of epistemology,* 168–187. Norwood, New Jersey: Ablex.

Gołąb, Zbigniew. 1964. *Conditionalis typu bałkańskiego w językach południowosłowiańskich, ze szczególnym uwzględnieniem macedońskiego* (Polska Akademia Nauk, Oddział w Krakowie, Prace Komisji Językoznawstwa 2). Wrocław – Kraków – Warszawa: Polska Akademia Nauk.

Gołąb, Zbigniew. 1990. The ethnic background and internal linguistic mechanism of the so-called Balkanization of Macedonian. *Balkanistica* 10. 13—19.

Granqvist, Kimmo. 2011. *Lyhyt Suomen romanikielen kielioppi.* [A short grammar of Finnish Romani.] (Kotimaisten kielten tutkimuskeskuksen verkkojulkaisuja 24.) Helsinki: Kotimaisten kielten tutkimuskeskus.

Graves, Nina. 2000. Macedonian – a language with three perfects? Dahl, Östen (ed.), *Tense and aspect in the languages of Europe* (Empirical Approaches to Language Typology 20-6 = Eurotyp vol. 6), 479–494. Berlin: Mouton de Gruyter.

Gumperz, John J. & Robert Wilson. 1971. Convergence and creolization: A case from the Indo-Aryan / Dravidian border in India. In Hymes, Dell (ed.), *Pidginization and creolization of languages*, 151–167. Cambridge: Cambridge University Press.

[Hauge, Kjetil Rå] Hauge, Hetil Ro. 1977. Sintaktični balkanizmi v bălgarskija ezik i ezikova redundantnost. *Bălgarski ezik* 27(5). 380–385.

Hinrichs, Uwe. 2004. Orale Kultur, Mehrsprachigkeit, radikaler Analytismus: Zur Erklärung von Sprachstrukturen auf dem Balkan und im kreolischen Raum. Ein Beitrag zur Entmystifizierung der Balkanlinguistik. *Zeitschrift für Balkanologie* 40(2). 141–174.

Horrocks, Geoffrey. 2010. *Greek: A history of the language and its speakers.* Oxford: Wiley-Blackwell.

Hristov, Petko. 2008. Mobilités du travail (gurbet), stratégies sociales et familiale: Une étude de cas dans les Balkans centraux. *Balkanologie* 11(1-2). 2–14.

Joseph, Brian D. 2013. Language contact in the Balkans. In Hickey, Raymond (ed.), *The handbook of language contact,* 618–633. Chichester: Blackwell.

Kusters, Wouter. 2003. *Linguistic complexity: The influence of social change on verbal inflection.* Utrecht: LOT. (Doctoral dissertation, University of Leiden.)

Lindstedt, Jouko. 2000a. Linguistic Balkanization: Contact-induced change by mutual reinforcement. In Dicky Gilbers, John Nerbonne & Jos Schaeken (eds.), *Languages in Contact* (Studies in Slavic and General Linguistics 28), 231–246. Amsterdam: Rodopi.

Lindstedt, Jouko. 2000b. The perfect — aspectual, temporal and evidential. Dahl, Östen (ed.), *Tense and aspect in the languages of Europe* (Empirical Approaches to Language Typology 20-6 = Eurotyp vol. 6), 365–383. Berlin: Mouton de Gruyter.

Lindstedt, Jouko. 2002. Is there a Balkan verb system? *Balkanistica* 15. 323–336.

McWhorter, John H. 2001. The world's simplest grammars are creole grammars. *Linguistic typology* 5(2/3). 125–166.

McWhorter, John H. 2011. *Linguistic simplicity and complexity: Why do languages undress?* (Language Contact and Bilingualism 1.) Boston: De Gruyter Mouton.

Markovik', Marjan. 2007. *Aromanskiot i makedonskiot govor od ohridsko-struškiot region vo balkanski kontekst.* Skopje: Makedonskata akademija na naukite i umetnostite.

[Makartsev, Maxim] Makarcev, M. M. 2014. *Èvidencial'nost' v prostranstve balkanskogo teksta.* Moskva – Sankt-Peterburg: Nestor-Istorija; Rossijskaja akademija nauk, Institut slavjanovedenija.

Sandfeld, Kristian. 1930. *Linguistique balkanique: Problèmes et résultats* (Collection linguistique publiée par la Société de Linguistique de Paris 31). Paris: Klincksieck. (Noveau tirage 1968.)

Thomason, Sarah G. 2001. *Language contact: An introduction.* Washington, D.C.: Georgetown University Press.

Trudgill, Peter. 2002. Linguistic and social typology. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change,* 707–728. Malden, MA: Blackwell.

Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity.* Oxford: Oxford University Press.

Wahlström, Max. 2015. *The loss of case inflection in Bulgarian and Macedonian* (Slavica Helsingiensia 47). Helsinki: University of Helsinki, Department of Modern Languages.

Weigand, Gustav. 1895. *Die Aromunen. Ethnographisch-philologisch-historische Untersuchungen über das Volk der sogenannten Makedo-Romanen oder Zinzaren.* Erster Band: Land und Leute. Leipzig: Johann Ambrosius Barth (Arthur Meiner).

Weinreich, Uriel. 1970. *Languages in contact: Findings and problems.* Seventh printing. The Hague: Mouton. (First edition in 1953.)